## REMARKS

Claims 1-4 are pending in the instant application. Claims 1-4 have been rejected. Claim 1 has been amended. Claims 2 and 4 have been canceled. No new matter has been added by this amendment. Reconsideration is respectfully requested in light of these amendments and the following remarks.

### I. Rejection of Claims Under 35 U.S.C. §103

Claims 1-4 have been rejected under 35 U.S.C. 103(a) as being unpatentable over Troyanskaya et al. ((2001) *Bioinformatics* 17:520-525) in view of Cunningham (US 2002/0129038) with additional support from online Merriam-Webster Dictionary ("Gaussian" definition). It is suggested that Troyanskaya et al. describe methods for estimating missing values in DNA microarrays via imputing, wherein k-means clustering and various model-based approaches and algorithms, such as (Single Value Decomposition) SVDimpute algorithm via normalization for microarray data comprising rows and columns, are described. It is suggested that according to the online Merriam-Webster dictionary, the definition of "Gaussian" is "being or having the shape of a normal curve or a normal distribution". It is further suggested that Troyanskaya et al. describe using k eigengenes, using a row average, and an expectation maximization method that is repeated until the change falls below a threshold. The Examiner acknowledges that Troyanskaya et al. do not recite a model which imposes a mixture of multivariate normal distributions; however, it is suggested that Cunningham describes a computer system and computer readable media

with data storage devices as well as an expectation-maximization algorithm that is performed in a computer implemented data mining system to create a Gaussian Mixture Model as well as generating output describing clustering in the data by computing a mixture of multivariate normal distributions.

Claims 1 and 3 have been rejected under 35 U.S.C. 103(a) as being unpatentable over Hytopoulos et al. (US 2002/0169560 A1) in view of Cunningham (US 2002/0129038) with additional support from online Merriam-Webster dictionary ("Gaussian" definition). It is suggested that Hytopoulos et al. describe a computer-implemented method and a system using microarray expression data arrays, cluster arrays, and clustering tools wherein the expression values have been normalized, filtered, and imputed, wherein missing data are imputed, and outputted. The Examiner further suggests that Hytopoulos et al. describe allowing the user to select K-nearest neighbor imputation mechanism or other data imputation mechanisms and the analysis of gene expression data to form clusters. IT is suggested that Hytopoulos et al. describe identifying genes represented in respective rows, which represents a partitioning of rows of microarray data, and mapping rows of expression data. The Examiner acknowledges that Hytopoulos et al. do not describe a model which imposes a mixture of multivariate normal distributions; however, Cunningham is suggested to describe a computer system and computer readable media with data storage devices as well as an expectation-maximization algorithm that is performed in a computer implemented data mining system to create a Gaussian Mixture Model as well as generating output describing clustering in the data by computing a mixture of multivariate normal distributions.

Claims 1-4 have been rejected under 35 U.S.C. 103(a) as being unpatentable over Hytopoulos et al. (US 2002/0169560 A1) with additional support from online Merriam-Webster dictionary ("Gaussian" definition) in view of Cereghini et al. (US 6,496,834 B1). It is suggested that Hytopoulos et al. describe the limitations in instant claims 1 and 3. The Examiner acknowledges that Hytopoulos et al. do not describe repeating a classification expectation-maximization algorithm until the K partitions converge or a model which imposes a mixture of multivariate normal distributions; however, it is suggested that Cereghini et al. describe a method of performing cluster analysis inside a relational database management system using Gaussian mixture parameters and implementing an Expectation-Maximization (EM) clustering algorithm iteratively. It is further suggested that Cereghini et al. describe grouping a set of data into k clusters with k rows (partitioned) and the expectation-maximization algorithm converges quickly and performing iterations. It is suggested that Cereghini et al. describe the EM algorithm assumes the data is formed by the mixture of multivariate normal distributions and is robust for noisy data and missing information such that the combination of these references would be obvious.

Applicants respectfully disagree with these rejections. In particular, it is respectfully submitted that while the Examiner has suggested that the cited documents purportedly teach the step of clustering the data by a Gaussian mixture clustering (GMC) model including an expectation maximization method that is repeated until the change falls below a threshold, the Examiner's rejections fail to identify with any particularity where the cited documents teach

the step of estimating missing values by a GMCimpute algorithm. As described in the paragraph spanning pages 17 and 18, the GMCimpute algorithm takes the average of all the K_estimates by *S* components, wherein each missing entry has *S* estimates and the final estimate is the average of them.

It is elementary that to support an obviousness rejection all words in a claim must be considered in judging the patentability of that claim against the prior art. *In re Wilson*, 424 F.2d 1382, 1385 (CCPA 1970).

In addition to failing to teach step (b) as presently set forth in claim 1, it is respectfully submitted that the cited documents also fail to teach or suggest the use of Bayesian information criterion to select the number of clusters *K* in the Gaussian mixture clustering as disclosed at paragraph spanning pages 13 and 14. Accordingly, in an earnest effort to further distinguish the present invention from the teachings of the cited references, Applicants have amended claim 1 to incorporate the subject matter of claim 2 and specify the use of Bayesian information criterion to select the number of clusters *K*. Support for this amendment is found at pages 13 and 14 and in claims 2 and 4. Accordingly, claims 2 and 4 have been canceled.

The framework for the objective analysis for determining obviousness under 35 U.S.C. 103 is stated in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966). Obviousness is a question of law based on underlying factual inquiries. The factual inquiries are as follows:

(A) Ascertaining the differences between the claimed invention and the prior art; and

(B) Ascertaining the differences between the claimed invention and the prior art; and

(C) Resolving the level of ordinary skill in the pertinent art. See MPEP 2141.

In the instant case, the combined teachings of the cited documents simply fail to teach each and every element of the claimed invention. Therefore, these references cannot be held to make the present invention obvious under 35 U.S.C. 103(a). It is therefore respectfully requested that these rejections be reconsidered and withdrawn.

**II. Conclusion**

Applicants believe that the foregoing comprises a full and complete response to the Office Action of record. Accordingly,

favorable reconsideration and subsequent allowance of the pending claims is earnestly solicited.

Respectfully submitted,

Jane Massey Licata
Registration No. 32,257

Date: **September 3, 2008**

Licata & Tyrrell P.C.
66 E. Main Street
Marlton, New Jersey 08053

(856) 810-1515